News Editorials: Towards Summarizing Long Argumentative Texts (COLING 2020)

Shahbaz Syed
Roxanne El Baff
Khalid Al-Khatib
Johannes Kiesel
Benno Stein
Martin Potthast
RAPE IS A VIOLENT CRIME.

At UNC and colleges across the country, it is treated as an infraction.

Something so fundamentally wrong demands a principled solution. This means undertaking a complete overhaul of the way the University deals with cases of rape. And it means recognizing for what it is: a systematic problem of institutional indifference, and instead creating a system that recognizes the University's responsibility to protect its students. A culture of rape justifies the use of a broken system, and should be replaced with a focus on:

1. Support for survivors;
2. Education for all students on the consequences of rape;
3. A defense in the pursuit of justice to trained law enforcement.

Words from the Watermark leadership

Jason Thomas, publisher of The News & Observer, which is a part of the News & Observer Media Group, was able to disprove the notion that the University is not taking the issue seriously. He emphasized the importance of educating students about the consequences of rape, and the need for support for survivors. He also highlighted the need for a defense in the pursuit of justice to trained law enforcement.

Todd Wagner: The goal of being consistent is to establish a personal relationship with the

By A.G. Noorani

Banning books

BOOK banning is akin to book burning, emblematic of Hitler's Germany. It assumes overwhelming importance in any discussion on intellectual freedom. The banning of books, as many would argue, is a form of censorship. It is a form of control that limits the flow of information, and can be used to suppress dissent and critical thinking.

In the context of India, recent years have witnessed a rise in the number of book burnings and book bans. The most notable example being the banning of the book 'The Satanic Verses' by Salman Rushdie in 1988. The Indian government, under the leadership of Prime Minister Rajiv Gandhi, banned the book in 1988, citing the threat of communal strife.

Such a ban on a book's sale is long overdue, but the challenge remains one that demands to be addressed. The banning of a book cannot solve the problem of a book being sold. It can only silence the voices of those who seek to read and discuss it.

In conclusion, the banning of books is a violation of the democratic process and a violation of the fundamental rights of citizens. It is a step towards the suppression of free speech and the stifling of intellectual freedom. It is a step towards the creation of a society that is not only controlled, but is also afraid of challenging the status quo.

In an interview with the author, "The Satanic Verses" author, Rushdie said, "I am happy that the book is not banned in India. It is a book that I would like to see banned in India. It is a book that I would like to read in India. It is a book that I would like to discuss in India. It is a book that I would like to be part of a conversation in India."

This new form of intolerance reflects the political and religious influence which the book has opened up to new territories. In January 2016, the Hindi University in New Delhi banned the book in the wake of protests against it.

A new form of censorship has emerged, targeting the foreign author.

An official of the central government, who is authorized to issue such a ban, said that the book was not适当的 for the public. The ban was enforced by the University's administration, which cited reasons of "national security" and "public interest".

This is not the first time that a book has been banned in India. In 1993, the book 'The Satanic Verses' by Rushdie was banned by several Indian states, including Gujarat and Madhya Pradesh.

In conclusion, the banning of books is a violation of the democratic process and a violation of the fundamental rights of citizens. It is a step towards the suppression of free speech and the stifling of intellectual freedom. It is a step towards the creation of a society that is not only controlled, but is also afraid of challenging the status quo.
News Reports vs Editorials

Reports

- Who? What? When?
- Where? Why? How?
- Details
- General & Background Information
- Importance

- **Goal**: Inform
- **Style**: Objective
- **Layout**: Inverted Pyramid

[Wikipedia]
News Reports vs Editorials

Reports

- **Goal**: Inform
- **Style**: Objective
- **Layout**: Inverted Pyramid

Editorials

- **Goal**: Persuade
- **Style**: Subjective
- **Layout**: Lead-Body-Conclusion

[Diagram showing the structure of reports and editorials]

- Details
- General & Background Information

[wikipedia link]
News Reports vs Editorials

**Reports**
- **Goal**: Inform
- **Style**: Objective
- **Layout**: Inverted Pyramid

**Editorials**
- **Goal**: Persuade
- **Style**: Subjective
- **Layout**: Lead-Body-Conclusion

---

[Introduces the issue]

[Wikipedia]
News Reports vs Editorials

Reports

- **Goal**: Inform
- **Style**: Objective
- **Layout**: Inverted Pyramid

Editorials

- **Goal**: Persuade
- **Style**: Subjective
- **Layout**: Lead-Body-Conclusion
News Reports vs Editorials

Reports

- **Goal**: Inform
- **Style**: Objective
- **Layout**: Inverted Pyramid

**Who? What? When?**
**Where? Why? How?**
**Details**
**General & Background Information**

Editorials

- **Goal**: Persuade
- **Style**: Subjective
- **Layout**: Lead-Body-Conclusion

**LEAD**
**BODY**
**CONCLUSION**

- Introduces the issue
- Provides arguments & Background
- Implicit suggestions / Call for action

[Wikipedia]
News Reports vs Editorials

**Reports**
- **Goal**: Inform
- **Style**: Objective
- **Layout**: Inverted Pyramid
- **Summary**: Top-n sentences (e.g. LEAD-3)

---

**Editorials**
- **Goal**: Persuade
- **Style**: Subjective
- **Layout**: Lead-Body-Conclusion

---

**Who? What? When?**
- Importance
  - Introduces the issue
  - Provides arguments & Background
  - Implicit suggestions / Call for action
News Reports vs Editorials

Reports

- **Goal**: Inform
- **Style**: Objective
- **Layout**: Inverted Pyramid
- **Summary**: Top-n sentences (e.g. LEAD-3)

Editorials

- **Goal**: Persuade
- **Style**: Subjective
- **Layout**: Lead-Body-Conclusion
- **Summary**: ?

[Wikipedia]
This Paper

Research Question:
What is a high-quality summary of an editorial?
This Paper

Research Question:
What is a high-quality summary of an editorial?

Hypothesis:
It preserves the core message and the most persuasive segments; it is structurally similar to the editorial
This Paper

Research Question:
What is a high-quality summary of an editorial?

Hypothesis:
It preserves the core message and the most persuasive segments; it is structurally similar to the editorial

Approach:
Define a high-quality summary; operationalize the annotation and evaluation of summaries
This Paper

Research Question:
What is a high-quality summary of an editorial?

Hypothesis:
It preserves the core message and the most persuasive segments; it is structurally similar to the editorial.

Approach:
Define a high-quality summary; operationalize the annotation and evaluation of summaries.

Contributions:
1. A unified scheme for annotation and evaluation of editorial summaries.
2. Webis-EditorialSum-2020 Corpus containing 1330 summaries for 266 news editorials.
3. Detailed analyses w.r.t summary structure, summary quality and its argumentation.
Anatomy of an Editorial

LEAD

BODY

CONCLUSION
Anatomy of an Editorial

Argumentative Discourse Units
Anatomy of an Editorial

Argumentative Discourse Units

Argumentative Structure
- Thesis
- Justification
Anatomy of an Editorial

- Argumentative Discourse Units
- Argumentative Structure
  - Thesis
  - Justification
- Evidence Types
  - Anecdote
  - Assumption
  - Common-Ground
  - Statistic
  - Testimony

A News Editorial Corpus for Mining Argumentation Strategies [Al-Khatib et al., 2016]
Evidence Types

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anecdote</td>
<td>We were in-between lessons during our first class, when we suddenly heard the sound of shooting.</td>
</tr>
</tbody>
</table>
# Evidence Types

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Anecdote</em></td>
<td>We were in-between lessons during our first class, when we suddenly heard the sound of shooting.</td>
</tr>
<tr>
<td><em>Assumption</em></td>
<td>Many have simply lost faith in global climate negotiation summits such as COP 20 starting in Lima, Peru, today.</td>
</tr>
</tbody>
</table>
## Evidence Types

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anecdote</td>
<td>We were in-between lessons during our first class, when we suddenly heard the sound of shooting.</td>
</tr>
<tr>
<td>Assumption</td>
<td>Many have simply lost faith in global climate negotiation summits such as COP 20 starting in Lima, Peru, today.</td>
</tr>
<tr>
<td>Common-Ground</td>
<td>Politicians are meant to act in the interests of their people.</td>
</tr>
</tbody>
</table>
### Evidence Types

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anecdote</td>
<td>We were in-between lessons during our first class, when we suddenly heard the sound of shooting.</td>
</tr>
<tr>
<td>Assumption</td>
<td>Many have simply lost faith in global climate negotiation summits such as COP 20 starting in Lima, Peru, today.</td>
</tr>
<tr>
<td>Common-Ground</td>
<td>Politicians are meant to act in the interests of their people.</td>
</tr>
<tr>
<td>Statistic</td>
<td>In the early 1900s, Argentina ranked among the world’s top 10 in per capita income.</td>
</tr>
</tbody>
</table>
Evidence Types

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anecdote</td>
<td>We were in-between lessons during our first class, when we suddenly heard the sound of shooting.</td>
</tr>
<tr>
<td>Assumption</td>
<td>Many have simply lost faith in global climate negotiation summits such as COP 20 starting in Lima, Peru, today.</td>
</tr>
<tr>
<td>Common-Ground</td>
<td>Politicians are meant to act in the interests of their people.</td>
</tr>
<tr>
<td>Statistic</td>
<td>In the early 1900s, Argentina ranked among the world’s top 10 in per capita income.</td>
</tr>
<tr>
<td>Testimony</td>
<td>“I saw my brother drown in front of my eyes,” said Hamid.</td>
</tr>
</tbody>
</table>
Summarizing an Editorial

Triangle: Thesis  ▲: Justification
Summary should preserve the thesis and the most persuasive segments (presuming extractive summarization)
Summarizing an Editorial

Summary

- Summary should preserve the **thesis** and the **most persuasive segments** (presuming extractive summarization)
- A concrete definition of a **high-quality** summary is lacking

Thesis  Justification
Summarizing an Editorial

Summary should preserve the **thesis** and the **most persuasive segments** (presuming extractive summarization).

A concrete definition of a **high-quality** summary is lacking.

Scholarship on **argument quality assessment** provides the foundation [Wachsmuth et al., 2017a, 2017b].

Thesis  Justification

Computational Argumentation Quality Assessment in Natural Language [Wachsmuth et al., 2017a]
Argumentation Quality Assessment: Theory vs. Practice [Wachsmuth et al., 2017b]
Quality Dimensions

1. **Thesis-indicativeness**
   Summary explicitly contains the thesis or indicate its
Quality Dimensions

1. **Thesis-indicativeness**
   Summary explicitly contains the thesis or indicate its

2. **Persuasiveness**
   Summary aims to persuade the target audience of its thesis
Quality Dimensions

1. **Thesis-indicativiveness**
   Summary explicitly contains the thesis or indicate its

2. **Persuasiveness**
   Summary aims to persuade the target audience of its thesis

3. **Reasonableness**
   Summary helps to reach the thesis and rebut plausible counter-arguments to it
Quality Dimensions

1. **Thesis-indicativeness**
   Summary explicitly contains the thesis or indicate its

2. **Persuasiveness**
   Summary aims to persuade the target audience of its thesis

3. **Reasonableness**
   Summary helps to reach the thesis and rebut plausible counter-arguments to it

4. **Conciseness**
   Summary is shorter than the editorial and avoid superfluous details
Quality Dimensions

1. **Thesis-indicativeness**
   Summary explicitly contains the thesis or indicate its

2. **Persuasiveness**
   Summary aims to persuade the target audience of its thesis

3. **Reasonableness**
   Summary helps to reach the thesis and rebut plausible counter-arguments to it

4. **Conciseness**
   Summary is shorter than the editorial and avoid superfluous details

5. **Self-containedness**
   Summary is comprehensible without referring to additional resources
Summary Creation
Summary Creation

Editorial

Summaries

⚠️ Thesis ⬅️ Justification
Summary Creation

- **Webis-Editorials-16 Corpus** of 300 editorials (each labeled with evidence types) as our data source [Al-Khatib et al. 2016]
- Excluded short editorials resulting in 266 editorials; summary length is fixed to 20%
- Each editorial was annotated by five workers with labels ▲ Thesis △ Justification
- Each summary can have up to two thesis segments

A News Editorial Corpus for Mining Argumentation Strategies [Al-Khatib et al., 2016]
Summary Evaluation

- Quality dimensions
  - Thesis-indicativeness, Persuasiveness, Reasonableness, Self-containedness
Summary Evaluation

- **Quality dimensions**
  - Thesis-indicativeness, Persuasiveness, Reasonableness, Self-containedness

- Each summary is evaluated by **three** workers; on a 4-point Likert scale
  (strongly-disagree, weakly-disagree, weakly-agree, strongly-agree)
Summary Evaluation

- **Quality dimensions**
  Thesis-indicativeness, Persuasiveness, Reasonableness, Self-containedness

- Each summary is evaluated by three workers; on a 4-point Likert scale
  (strongly-disagree, weakly-disagree, weakly-agree, strongly-agree)

- Only the title of the editorial is shown alongside each summary
Summary Evaluation

- **Quality dimensions**
  - Thesis-indicativeness, Persuasiveness, Reasonableness, Self-containedness

- Each summary is evaluated by **three workers**; on a 4-point Likert scale
  (strongly-disagree, weakly-disagree, weakly-agree, strongly-agree)

- Only the title of the editorial is shown alongside each summary

- A high-quality summary has at least three dimensions (according to majority vote)
Key Findings

- 73.2% Thesis are from the LEAD while 67.6% of Justification are from the BODY
Key Findings

- 73.2% Thesis are from the LEAD while 67.6% of Justification are from the BODY
- Summaries adhere to the three-part layout of editorials
  - LEAD (28.3%), BODY (60.6%), CONCLUSION (11.1%)
Key Findings

- 73.2% Thesis are from the LEAD while 67.6% of Justification are from the BODY
- Summaries adhere to the three-part layout of editorials
  - LEAD (28.3%), BODY (60.6%), CONCLUSION (11.1%)
- High-quality summaries have more third-party evidence (Statistic and Testimony)
Key Findings

- 73.2% Thesis are from the LEAD while 67.6% of Justification are from the BODY
- Summaries adhere to the three-part layout of editorials:
  - LEAD (28.3%), BODY (60.6%), CONCLUSION (11.1%)
- High-quality summaries have more third-party evidence (Statistic and Testimony)
- Reasonableness is significantly correlated with Persuasiveness and Self-Containedness
Example

2015: Beyond Obama, new Congress, we need a revival of the American spirit.  [Fox News, 2015]

High-quality summary

A Pew Research Center study confirms one trend that has been obvious over several years. The "typical" American family is no longer typical. Just 46 percent of American children now live in homes with their married, heterosexual parents. Five percent have no parents at home. They most likely are living with grandparents, says the study. A major contributor to this trend has been the assault on marriage and other institutions by the Baby Boom generation. It was that generation that promoted cohabitation, no-fault divorce, hatred of the police (they called them "pigs" then, too) and disdain for the military and America, spawned not just by the Vietnam War but a life of relative ease unknown to their parents.

<table>
<thead>
<tr>
<th>Majority</th>
<th>Thesis-indicativeness</th>
<th>Persuasiveness</th>
<th>Reasonableness</th>
<th>Self-containedness</th>
</tr>
</thead>
<tbody>
<tr>
<td>✔️</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Example

2015: Beyond Obama, new Congress, we need a revival of the American spirit. [Fox News, 2015]

Low-quality summary

the stock market is at record highs. A Pew Research Center study confirms one trend that has been obvious over several years. In the film, "Girl Interrupted," Winona Ryder plays an 18-year-old who enters a mental institution for what is diagnosed as borderline personality disorder. These startling figures about the decline of the American family contrast with the year 1960. It was that generation that promoted cohabitation, no-fault divorce, hatred of the police (they called them "pigs" then, too) and disdain for the military and America, spawned not just by the Vietnam War but a life of relative ease unknown to their parents. Since the elimination of the draft, young people are no longer expected to serve in the military and so most of them pursue whatever goals they wish with no expectation they should give something back to their country. All empires begin, writes Glubb, with the age of pioneers, followed by ages of conquest, commerce, affluence, intellect and decadence.

<table>
<thead>
<tr>
<th>Majority</th>
<th>Thesis-indicativeness</th>
<th>Persuasiveness</th>
<th>Reasonableness</th>
<th>Self-containedness</th>
</tr>
</thead>
<tbody>
<tr>
<td>❌</td>
<td>❌</td>
<td>❌</td>
<td>❌</td>
<td>❌</td>
</tr>
</tbody>
</table>
Conclusion

Contributions

- A unified scheme for annotating and evaluating summaries of editorials
- **Webis-EditorialSum-2020 Corpus** containing 1330 summaries for 266 editorials
  (90% of editorials have at least 3 high-quality summaries, 52% have 5 high-quality summaries)
- Detailed analyses of summary structure, evidence types and summary quality
Conclusion

Contributions

- A unified scheme for annotating and evaluating summaries of editorials
- **Webis-EditorialSum-2020 Corpus** containing 1330 summaries for 266 editorials  
  (90% of editorials have at least 3 high-quality summaries, 52% have 5 high-quality summaries)
- Detailed analyses of summary structure, evidence types and summary quality

Discussion

- Summaries are extractive in nature
- Larger corpora following our annotation scheme are required for training end-to-end models
Conclusion

Contributions

- A unified scheme for annotating and evaluating summaries of editorials
- **Webis-EditorialSum-2020 Corpus** containing 1330 summaries for 266 editorials
  (90% of editorials have at least 3 high-quality summaries, 52% have 5 high-quality summaries)
- Detailed analyses of summary structure, evidence types and summary quality

Discussion

- Summaries are extractive in nature
- Larger corpora following our annotation scheme are required for training end-to-end models

<table>
<thead>
<tr>
<th>Corpus</th>
<th>Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="webis.de/data/webis-editorialsum-20">webis.de/data/webis-editorialsum-20</a></td>
<td><a href="github.com/webis-de/coling-20">github.com/webis-de/coling-20</a></td>
</tr>
</tbody>
</table>